1 Model

The model takes the form

$$\boldsymbol{x}_k = \boldsymbol{f}(\boldsymbol{x}_{k-1}) + \boldsymbol{v}_k \tag{1a}$$

$$\boldsymbol{y}_k = \boldsymbol{H}\boldsymbol{x}_{k-1} + \boldsymbol{w}_k \tag{1b}$$

where $\boldsymbol{v}_k$ and $\boldsymbol{w}_k$ are white, uncorrelated, zero mean Gaussian noise vectors with covariances of $\boldsymbol{Q}$ and $\boldsymbol{R}$, respectively. For this model the transition density is defined as

$$p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}) = \mathcal{N}(\boldsymbol{f}(\boldsymbol{x}_{k-1})), \boldsymbol{Q}) \tag{2}$$

$$= (2\pi)^{-d/2}|\boldsymbol{Q}|^{-1/2} \exp\left(-\frac{1}{2}(\boldsymbol{x}_k - \boldsymbol{f}(\boldsymbol{x}_{k-1}))^T \boldsymbol{Q}^{-1}(\boldsymbol{x}_k - \boldsymbol{f}(\boldsymbol{x}_{k-1}))\right) \tag{3}$$

and the likelihood is

$$p(\boldsymbol{y}_k|\boldsymbol{x}_k) = \mathcal{N}(\boldsymbol{H}\boldsymbol{x}_k, \boldsymbol{R}) \tag{4}$$

$$= (2\pi)^{-n/2}|\boldsymbol{R}|^{-1/2} \exp\left(-\frac{1}{2}(\boldsymbol{y}_k - \boldsymbol{H}\boldsymbol{x}_k)^T \boldsymbol{R}^{-1}(\boldsymbol{y}_k - \boldsymbol{H}\boldsymbol{x}_k)\right) \tag{5}$$

2 PCRB

The PCRB is computed from the Fisher information matrix computed on the joint density $p(\boldsymbol{y}, \boldsymbol{x})$,

$$\boldsymbol{J} = \mathbb{E}\left[\nabla_{\boldsymbol{x}}\nabla_{\boldsymbol{x}}^T \log p(\boldsymbol{y}, \boldsymbol{x})\right] \tag{6}$$

where the expectation is taken over $\boldsymbol{y}$ and $\boldsymbol{x}$. The PCRB bounds the covariance of an unbiased estimator, $\boldsymbol{P}$,

$$\boldsymbol{P} \geq \boldsymbol{J}^{-1} \tag{7}$$

Recursive formulas are based on the logarithm of the likelihood and transition densities, which for the model above are

$$\log p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}) = C_1 - \frac{1}{2}(\boldsymbol{x}_k - \boldsymbol{f}(\boldsymbol{x}_{k-1}))^T \boldsymbol{Q}^{-1}(\boldsymbol{x}_k - \boldsymbol{f}(\boldsymbol{x}_{k-1})) \tag{8}$$

$$\log p(\boldsymbol{y}_k|\boldsymbol{x}_k) = C_2 - \frac{1}{2}(\boldsymbol{y}_k - \boldsymbol{H}\boldsymbol{x}_k)^T \boldsymbol{R}^{-1}(\boldsymbol{y}_k - \boldsymbol{H}\boldsymbol{x}_k) \tag{9}$$

2.1 Fisher information recursion (non-singular Q)

A recursive formula for the FIM, $\boldsymbol{J}_k$, was developed in [1]. The recursion proceeds as

$$\boldsymbol{J}_k(\boldsymbol{x}_k) = \boldsymbol{W}_k - \boldsymbol{V}_k^T \left[\boldsymbol{J}_{k-1}(\boldsymbol{x}_{k-1}) + \boldsymbol{U}_k\right]^{-1} \boldsymbol{V}_k \tag{10}$$

where

$$\boldsymbol{U}_k = -\mathbb{E}\left[\nabla_{\boldsymbol{x}_{k-1}}\nabla^T_{\boldsymbol{x}_{k-1}} \log p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})\right] \tag{11}$$

$$\boldsymbol{V}_k = -\mathbb{E}\left[\nabla_{\boldsymbol{x}_{k-1}}\nabla^T_{\boldsymbol{x}_{k}} \log p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})\right] \tag{12}$$

$$\boldsymbol{W}_k = -\mathbb{E}\left[\nabla_{\boldsymbol{x}_{k}}\nabla^T_{\boldsymbol{x}_{k}} \log p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})\right] - \mathbb{E}\left[\nabla_{\boldsymbol{x}_{k}}\nabla^T_{\boldsymbol{x}_{k}} \log p(\boldsymbol{y}_k|\boldsymbol{x}_{k})\right] \tag{13}$$

The derivatives for the above model are

$$\nabla^T_{\boldsymbol{x}_k} \log p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}) = -(\boldsymbol{x}_k - \boldsymbol{f}(\boldsymbol{x}_{k-1}))^T \boldsymbol{Q}^{-1} \tag{14}$$

$$\nabla_{\boldsymbol{x}_k}\nabla^T_{\boldsymbol{x}_k} \log p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}) = -\boldsymbol{Q}^{-1} \tag{15}$$

$$\nabla_{\boldsymbol{x}_{k-1}}\nabla^T_{\boldsymbol{x}_k} \log p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}) = \boldsymbol{F}(\boldsymbol{x}_{k-1})\boldsymbol{Q}^{-1} \tag{16}$$

$$\nabla^T_{\boldsymbol{x}_{k-1}} \log p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}) = (\boldsymbol{x}_k - \boldsymbol{f}(\boldsymbol{x}_{k-1}))^T \boldsymbol{Q}^{-1}\boldsymbol{F}(\boldsymbol{x}_{k-1}) \tag{17}$$

$$\nabla_{\boldsymbol{x}_{k-1}}\nabla^T_{\boldsymbol{x}_{k-1}} \log p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}) = -\boldsymbol{F}(\boldsymbol{x}_{k-1})^T\boldsymbol{Q}^{-1}\boldsymbol{F}(\boldsymbol{x}_{k-1}) \tag{18}$$

$$\nabla^T_{\boldsymbol{x}_k} \log p(\boldsymbol{y}_k|\boldsymbol{x}_k) = (\boldsymbol{y}_k - \boldsymbol{H}\boldsymbol{x}_k)^T \boldsymbol{R}^{-1}\boldsymbol{H} \tag{19}$$

$$\nabla_{\boldsymbol{x}_k}\nabla^T_{\boldsymbol{x}_k} \log p(\boldsymbol{y}_k|\boldsymbol{x}_k) = -\boldsymbol{H}^T \boldsymbol{R}^{-1}\boldsymbol{H} \tag{20}$$

Substituting these into (11)–(13) gives,

$$\boldsymbol{U}_k = \mathbb{E}\left[\boldsymbol{F}(\boldsymbol{x}_{k-1})^T\boldsymbol{Q}^{-1}\boldsymbol{F}(\boldsymbol{x}_{k-1})\right] \tag{21}$$

$$\boldsymbol{V}_k = \mathbb{E}\left[-\boldsymbol{F}(\boldsymbol{x}_{k-1})^T\boldsymbol{Q}^{-1}\right] \tag{22}$$

$$\boldsymbol{W}_k = \boldsymbol{H}^T\boldsymbol{R}^{-1}\boldsymbol{H} + \boldsymbol{Q}^{-1} \tag{23}$$

When noise is only added to a subset of the states each transition, $\boldsymbol{Q}$ is singular and $\boldsymbol{Q}^{-1}$ does not exist. Another method is required.

2.2 Cramer-Rao bound recursion (singular Q)

An alternative formulation is to recursively compute the bound at each step. That is, compute $\boldsymbol{P}_k = \boldsymbol{J}_k^{-1}$ directly rather than $\boldsymbol{J}_k$. The formulas suggested in [2] are applicable to wider class of models (e.g. multiplicative noise), denoted by a general transition function $\tilde{\boldsymbol{f}}(\boldsymbol{x}_k, \boldsymbol{v}_k)$. The recursion is

$$\boldsymbol{P}_k = \bar{\boldsymbol{F}}_{k-1}\left(\boldsymbol{P}^{-1}_{k-1} + \bar{\boldsymbol{R}}^{-1}_{k-1}\right)^{-1}\bar{\boldsymbol{F}}_{k-1} + \bar{\boldsymbol{G}}_{k-1}\boldsymbol{Q}_k\bar{\boldsymbol{G}}^{-1}_{k-1} \tag{24}$$

where

$$\bar{\boldsymbol{F}}_k = \mathbb{E}\left[\nabla^T_{\boldsymbol{x}_k}\tilde{\boldsymbol{f}}_k(\boldsymbol{x}_k, \boldsymbol{v}_k)\right] \tag{25}$$

$$\bar{\boldsymbol{R}}^{-1}_k = -\mathbb{E}\left[\nabla_{\boldsymbol{x}_k}\nabla^T_{\boldsymbol{x}_k} \log p(\boldsymbol{y}_k|\boldsymbol{x}_k)\right] \tag{26}$$

$$\bar{\boldsymbol{G}}_k = \mathbb{E}\left[\nabla^T_{\boldsymbol{v}_k}\tilde{\boldsymbol{f}}_k(\boldsymbol{x}_k, \boldsymbol{v}_k)\right] \tag{27}$$

$$\boldsymbol{Q}^{-1}_k = -\mathbb{E}\left[\nabla_{\boldsymbol{v}_k}\nabla^T_{\boldsymbol{v}_k} \log p(\boldsymbol{v}_k)\right] \tag{28}$$

and $\boldsymbol{P}_0^{-1}$ is the inverse of the prior covariance. The model in (1) is represented by $\tilde{\boldsymbol{f}}(\boldsymbol{x}_k, \boldsymbol{v}_k) = \boldsymbol{f}(\boldsymbol{x}_k) + \boldsymbol{v}_k$. This simplifies the derivatives to

$$\nabla^T_{\boldsymbol{x}_k} \tilde{\boldsymbol{f}}_k(\boldsymbol{x}_k, \boldsymbol{v}_k) = \boldsymbol{F}_k(\boldsymbol{x}_k) \tag{29}$$

$$\nabla^T_{\boldsymbol{v}_k} \tilde{\boldsymbol{f}}_k(\boldsymbol{x}_k, \boldsymbol{v}_k) = \boldsymbol{I}_{d\times d} \tag{30}$$

$$\nabla_{\boldsymbol{x}_k} \nabla^T_{\boldsymbol{x}_k} \log p(\boldsymbol{y}_k|\boldsymbol{x}_k) = -\boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{H} \tag{31}$$

and the corresponding matrices to

$$\bar{\boldsymbol{F}}_k = \mathbb{E}\left[\boldsymbol{F}_k(\boldsymbol{x}_k)\right] \tag{32}$$

$$\bar{\boldsymbol{R}}_k^{-1} = -\boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{H} \tag{33}$$

$$\bar{\boldsymbol{G}}_k = \boldsymbol{I}_{d\times d} \tag{34}$$

Notice that $\boldsymbol{Q}^{-1}$ is not directly required in the recursion, so this method is applicable to neural field models where process noise is only added to a single state.

2.3 Monte Carlo approximations

The expectations in (21) and (22) or (32) can be approximated by numerical integration using many trajectory realisations. That is,

$$\boldsymbol{U}_k \approx \frac{1}{M} \sum_{i=1}^{M} \boldsymbol{F}(\boldsymbol{x}_{k-1}^{(i)})^T \boldsymbol{Q}^{-1} \boldsymbol{F}(\boldsymbol{x}_{k-1}^{(i)}) \tag{35}$$

$$\boldsymbol{V}_k \approx \frac{1}{M} \sum_{i=1}^{M} \boldsymbol{F}(\boldsymbol{x}_{k-1}^{(i)})^T \boldsymbol{Q}^{-1} \tag{36}$$

or alternatively,

$$\bar{\boldsymbol{F}}_{k-1} \approx \frac{1}{M} \sum_{i=1}^{M} \boldsymbol{F}(\boldsymbol{x}_{k-1}^{(i)}) \tag{37}$$

3 Pendulum example

Let $\boldsymbol{x}_k = [a_k \quad b_k]^T$ where a_k is the pendulum angle and b_k is the angular velocity,

$$\boldsymbol{f}(\boldsymbol{x}_k) = \begin{bmatrix} a_k + b_k\Delta \\ -b_k - g\sin(a_k)\Delta \end{bmatrix} \tag{38}$$

The linearisation used for the EKF is

$$\boldsymbol{F}(\boldsymbol{x}_k) = \begin{bmatrix} 1 & \Delta \\ -g\cos(a_k) & 1 \end{bmatrix} \tag{39}$$

The observation matrix is $H = [1 \quad 0]$, so that only the first state element is measured.

3.1 Simulations

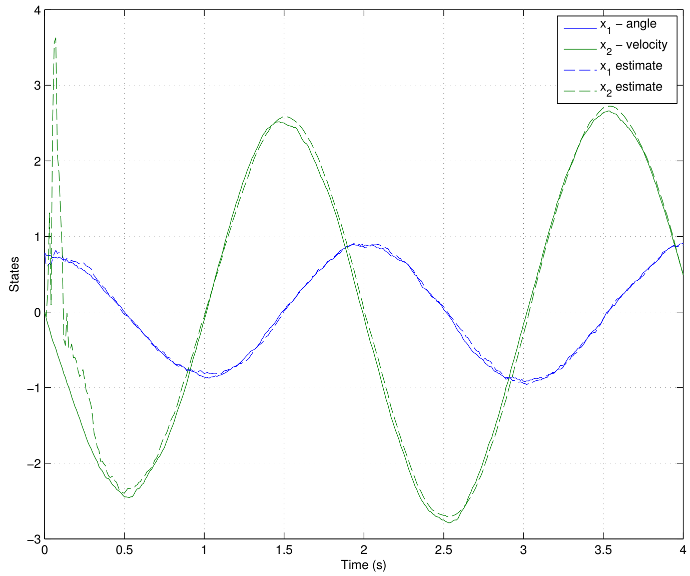

Figure 1: Trajectory of a pendulum and EKF estimates

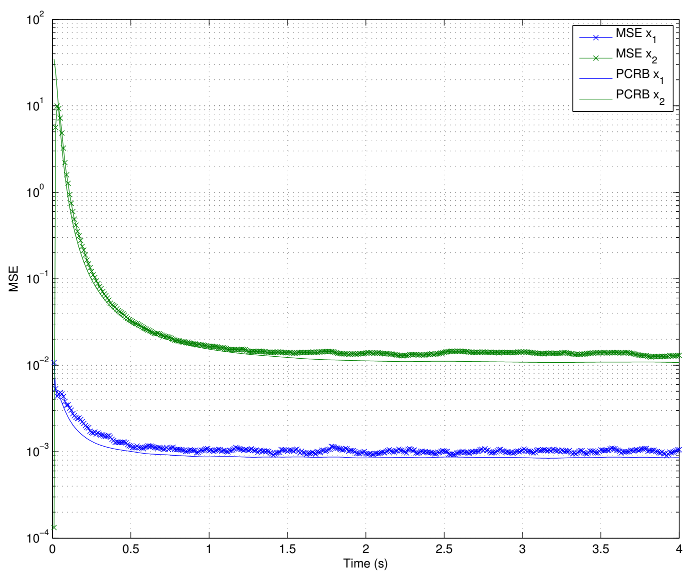

Figure 2: MSE of EKF estimates (averaged over 1000 realisations) compared to the PCRB. The suboptimal EKF is slightly above the PCRB.

References

[1] P. Tichavsky, C. H. Muravchik, and A. Nehorai, "Posterior Cramer-Rao bounds for discrete-time nonlinear filtering," *Signal Processing, IEEE Transactions on*, vol. 46, pp. 1386–1396, May 1998.

[2] N. Bergman, "Posterior Cramér-Rao bounds for sequential estimation," 2001.